EasyChair Preprint

№ 792

Categorization: A source of theory and output of research

Fred Niederman and Roman Lukyanenko

February 22, 2019

Categorization: A source of theory and output of research

Full Paper

Fred Niederman
Saint Louis University
3674 Lindell Blvd, 404
St. Louis, MO 63108
fred.niederman@slu.edu

Roman Lukyanenko
HEC Montréal
3000 Chemin de la Côte Sainte Catherine
Montréal, Québec H3T 2A7
roman.lukyanenko@hec.ca

ABSTRACT

In a research community, the use of the concept of category and categorization is widespread, generally helpful, but sometimes overly constraining. Despite the wealth of studies that propose new categories, a somewhat static view of categories pervades many disciplines. As we demonstrate on the analysis of a seminal framework by Gregor (2006), a given set of categories can be criticized and challenged in light of potentially valid alternatives. In contrast, we suggest for researchers to adopt the assumption of fluidity of categories, which leads to a different approach to demonstrating the contribution of research that deals with categories.

CCS CONCEPTS

• Social and professional topics~Management of computing and information systems

KEYWORDS

Classification, categorization, theory, research method

ACM Reference format:

F. Niederman, and R. Lukyanenko. 2018. Categorization: A source of theory and output of research. In *Proceedings of ACM SIGMIS Conference on Computers and People Research, Buffalo, New York USA, June 18-20, 2018*. 6 pages.

INTRODUCTION

> "[o]n those remote pages it is written that animals are divided into (a) those that belong to the Emperor, (b) embalmed ones, (c) those that are trained, (d) suckling pigs, (e) mermaids, (f) fabulous ones, (g) stray dogs, (h) those that are included in this classification, (i) those that tremble as if they were mad, (j) innumerable ones, (k) those drawn with a very fine camel's hair brush, (1) others, (m) those that have just broken a flower vase, (n) those that resemble flies from a distance" (from ancient Chinese encyclopedia 'Celestial Emporium of Benevolent Knowledge') [6].

> "[E]xisting typologies are likely to contain inconsistencies, trade-offs, and— perhaps most importantly—irrelevant elements. However, if not all parts of a configuration are equally important, the issue becomes this: Which are the critical aspects in a typology, and which elements are nonessential? The challenge of typologies thus is determining what really matters (and to what degree) in understanding the causal structure of a type." [14].

In a research community, the use of the concept of category [1] and categorization is widespread, generally helpful, but sometimes overly constraining. Recent management literature includes a stream of studies that focus on the actions and processes of "categorizing" rather than on static categories themselves [11, 14]. We believe, underlying the emergence of this stream is the fundamental difference in understanding of categories as fixed and determined (what Barsalou [1] calls "common categories") or whether they are fluid and explicitly constructed (what Barsalou [1] calls "ad hoc categories"), see also [26]. These differences correspond to worldviews that focus on objects and their permanence versus activities and how

[1] Regarding terminology, several words, categorization, typology, and taxonomy tend to be used interchangeably in some cases or with distinctions in others. Doty and Glick [11] make a clear but idiosyncratic set of distinctions among the terms. At times one or another term is used pejoratively to suggest something lesser than "theory" at other times rather neutrally. Our preference would be to refer to categorization as what Barsalou calls "ad hoc", typology, and taxonomy as common following the guidelines of static or "common category" types.

conceptualizations of objects by humans are in constant change.[2]

On the one hand understanding the concepts of category and categorization is a rather philosophical and linguistic endeavor, but it also has concrete implications regarding how we interpret collected knowledge. In this paper we will discuss at length the categorization of theory following Gregor [15] – a seminal categorization of theories that may exist in the information systems discipline. At the time of writing the paper was cited more than 2,300 times and serves as a significant reference for IS scholars. The intent is to consider how our interpretation of theory can change with the reconsideration of how to think about the nature of categories. It is our contention that a static view of categories and categorization will lead to different understandings about the nature of the discussion presented as well as our ability to extend them. In an ironic twist, one of the five categories of theory discussed by Gregor (2006) pertains to categorization as a type of theory. In this paper, we will discuss categorization per se using Gregor's typology as an example and further probe her first theory category exploring how categories can *BE* theory. It depends on which scholar one reads as to whether categorization is theory, implies theory, or sets the stage for further refinement [37, 48]. Not all scholars are completely consistent in their discussions of the nature of categories, which produces further fuzziness in thinking on this topic.

The thrust of our study presented below is to examine the nature of categories and categorization per se. We suggest a more nuanced view of the process by which a set of categories may be proposed and a more detailed description of what a category "is" within a set of categories. In other words, what does it mean to have such a set of categories as the one Gregor sets up? Although Barsalou [1] differentiates between common categories such as birds (p. 211) and "ad hoc" (e.g., things to take on a vacation, items in a shopping cart) - being categories that are constructed for a set purpose but may be fleeting. Yet, this is not a distinction that guided Gregor's (2006) work, which leaves open the potential for interpreting variations in how the typology might be viewed and applied.

In this paper, we will also illustrate the potential value of examining categorization as a serious enterprise, whether it is, leads to, or holds implicit theory using traditional IS personnel and inclusion issues.

[2] Note that objects in this case is a convenient reference to the whole array of entities including the physical and the socially constructed (e.g. a "table" and a "property deed")[4, 28, 45].

2 NATURE OF CATEGORIES

Much research in IS adopts what we call a "static view" of categories. A static view (sometimes referred to as "classical model") holds that they are permanent, meaningfully differentiated, and mutually exclusive (in that instances to be categorized can be determined to each belong to its underlying category [9, 31]). For example, when a bachelor is defined as "an unmarried man". Any instance (object) that shares these two attributes can be necessarily categorized as a bachelor.

Under this view, a category would simultaneously be describable as a set of rules whereby each new instance can be unambiguously sorted into the right group and where all members of each group share a unique set of attributes relative to the others. In other words, there should be both deductive and inductive access to each category (known in philosophy as intension vs extension, see [47]). If I see a new bird, I should be able to apply some combination of its attributes to determine if it is in the finch family or the jay family [37, 38]. At the same time, I should be able to gather say 100 finches and 100 jays and group them accordingly so that all members of each group share common attributes which are not shared with the other group (note that they may share values on one dimension but not all dimension – dogs and cats share mammalian values, but not common reproduction, for example). As Gregor (2006) implicitly points out this mapping of attributes to groups dates back to the thinking of Aristotle about the nature of things.

In contrast, a fluid or ad hoc view of categories would build categorization on the fly (pun intended). A group of 100 birds, whether finches or jays, might be grouped into those that are fast versus slow (for the purpose of gathering them by a fast worker versus a slow worker), large or small (for purposes of assigning them to cages), or hungry or already fed (for purposes of knowing which to feed next). Applying this to an organizational example, employees might be grouped into blue collar and white collar (for wage versus salary), vested and non-vested (for distribution of investment newsletters), smokers and non-smokers (for differentiating health care contributions) etc. Note that in both the bird and worker examples, such categorization is fluid (individuals move from unfed to fed and back repeatedly rather than staying in one category) and the instances may be difficult to assign (is a first level manager who does 80% of her work on the assembly line blue or white collar?) Note particularly how the categorization itself

is purposeful – if we pay the same for health care for smokers and non-smokers, there may no reason to create such a categorization and no one would bother [see also, 34]. It is also fluid as individuals shift between categories over time. Consider that when done with a threshold of care, it can be useful – why pay the extra smoking cost premium for health care for 100% of workers if only 20% are smokers?

The fact that objects do not inherently belong to categories, and that classification schemas are constant in a flux, has been argued repeatedly by psychologists, linguists, philosophers, natural and social scientists [5, 7, 19, 23, 31, 41]. Information systems researchers also begun to propose solutions that are predicated on a more dynamic view of categories [24, 36, 38]. Thus, Parsons and Wand [35], state: "[m]ultiple class structures can be constructed to model a domain of phenomena. Different structures may be useful for different purposes, and there is no inherently "correct" way of classifying the phenomena in a domain".

In information systems, the argument in support of this position is rooted in ontology – a branch of philosophy that studies what exists in reality [13, 25, 27, 36]. Specifically, "substantialist ontology" [12, 39], such as that of Mario Bunge, states that individuals or things are fundamental elements of existence [8]. The things possess properties (or, from the cognitive perspective, attributes). Classes are then formed based on some purpose by focusing on certain properties of interest, and ignoring others [22, 27, 33]. Thus, classification is an inherently purpose-driven, and therefore, always dynamic and fleeting process.

Consistent with this general perspective is the works of Lawrence Barsalou that extends these arguments in an important direction [1, 2, 31]. From this body of work, it is evident that notwithstanding the general ontological premise, some categories are more stable than others. These Barsalou [1] calls "common categories", and they are typically the categories that describe nature. As psychologists argue, in nature, attributes of things, correlate in bundles [43]. Thus, organisms that fly tend to have wings, organisms that have wings, tend to have low body metabolism, and organisms that have wings, tend to spend considerable time in the air and on high objects (e.g., flying, building nests). This fundamental aspect of reality results in an important function of categories – inferential utility [34, 38, 43]. Inferential utility means that when observing an organism, humans do not need to observe directly every property of this organism. Instead, many of these properties can be inferred from the observable ones. This makes classification fundamentally useful for humans. Building on this argument, researchers, such as Parsons and Wand [34, 38], have argued that sciences need to develop those categorization schemas – they called "classes" that follow the same natural process, and identify those categories that have inferential utility.

However, Barsalou argues that the inferential utility is not universal for categories. In particular, there is an important group of categories, those that are explicitly constructed (what Barsalou [1] calls "ad hoc categories"), see also [26]), which by design do not have many inferences. For example, we may form a category "things to take on vacation", which would contain many dissimilar objects that share only a single goal-oriented property. This "temporary" category would not have the same inferential utility as categories such as "bird" or "tree". That is, having observed some features of a member (e.g., that the passport is made of paper) of this category, it would be difficult to reliably infer additional features that would still be true for all members of this category (i.e., binoculars, walking sticks, swimming goggles, sun lotion, would not be made of paper). Nonetheless, these categories remain useful, but not for understanding reality, rather for communicating goals, perspectives, values as well as for effective social interaction and action. For example, we can use the single unifying property (things that are useful for the vacation), to scan our house for items fitting this criterion, and thus consider putting them in a suitcase.

In summary, based on the arguments in philosophy and psychology, we can conclude that a given object can be a member of multiple categories, that categories are constructed for a particular purpose. At the same time, some categories are more stable than others. Extrapolating these arguments into the task of creating categories to make a research contribution, we argue the following:

When researchers seek to devise a new categorization scheme, they may face two different scenarios. First, if the aim is to uncover some fundamental regularities in nature, they should seek to develop what Barsalou coined "common categories" or Parsons and Wand called "classes". While these categories would not be definitive and only categories applicable to the phenomena of interest, they should carve the domain at the important areas of discontinuities. As these categories neatly partition reality, they should have as little overlap as possible, and ideally none. For example, such natural categories as trees, birds and snakes (also called "basic-level categories) have very few common, shared attributes [10, 21, 26, 42]. In other words, these categories should exhibit strong internal cohesion, and very loose coupling.

The "common categories" should offer rich inferences – that is be made of inter-correlated and interdependent attributes, such that having identified some phenomena as a member of this category, one could infer additional, unobservable properties. Such categories would be more stable, and more long lasting, as long as the underlying phenomena does not change.

Second, if on the other hand, researchers seek to develop categories that are designed for effective communication about some scientific phenomena and for some specific action (e.g., compartmentalizing research papers into conference tracks), they should instead develop Barsalou's "ad hoc categories". These categories do not have to uncover fundamental regularities in the domain, and their utility is tied more directly to the usefulness and value of the goal behind these categories.

Based on the works of Barsalou [1], it appears that that the possibilities for additional dimensions by which theories may be differentiated are uncountable. As Barsalou shows, when driven by some purpose or utility, one can hone in on any attribute or feature of interest and arrive at completely novel ways to carve up reality. Additionally, this process underlies much of human experience, especially in social domains, where the utilitarian nature of categories is especially pronounced. For example, can we divide theory based on their amount of support (e.g. number and survival of tests) – if we did this I think we'd find that a couple of IS theories TAM and its descendants and McLean's success model have been well tested, but not many more in IS [3, 20, 46]. Can we divide theory based on ontological and/or epistemological assumptions or on complexity and number of elements? On types of relationships (e.g. that something can happen, that something usually happens, that something precedes something else; that two things correlate – as well as different kinds of causal forces)? Each of these distinctions can be meaningful for a particular application or purpose [27].

3 EXEMPLAR RESEARCH: GREGOR (2006)

We now apply our arguments to a prominent typology in the information systems discipline, "The nature of theory in information systems" by Shirley Gregor [15].

Gregor's (2006) typology of theory is presented within the static tradition. We are not critical of this whether the choice resulted from simply assuming that the static tradition represents the totality of conceptualizing typology or if it was a conscious decision. There are advantages of a static tradition such as the definitiveness of the categories, the easy recognition of the typology as a style per se, and of the vision of the categories as relatively permanent. However, more than a decade later, viewing the categorization framework from a fluid or ad hoc categorical perspective may suggest helpful new insights and extensions. Re-examining the fundamental assumptions is further timely, as both psychology, and, information systems studies that reference psychology, begin to recognize the value of a more nuanced approach to the nature of categories [26, 30].

The ad hoc perspective raises a number of questions not addressed by the original typology. How are the categories created (addressed briefly in Gregor (2006)? How do instances enter or exit from categories? How are the categories used? Note we do not ask are these the best categories (e.g. the best way to divide up the domain of instances into sub groups) but assume that it is one of many ways. As described by [38] these sorts of questions pertaining to categories open up new avenues of investigation in management broadly and in information systems research as well.

Actually Gregor uses one dimension of "goal" as a basis for developing categories whereas, she points out, Markus and Robey [29] present three dimensions each of which alone or in combination could be used to generate an alternative categorization scheme for theory. These proposed dimensions are: "…the nature of the causal agency (technology, organizational, or emergent); (2) the logical structure (whether *variance* or *process* theory); and (3) the level of analysis. Gregor 2006, p. 621)". It is clear that if the terrain is divided by one, two or all three of these, we will get different groupings of instances than we would derive from Gregor (2006).

Each of these three dimensions upon fine examination have their own limitations and challenges. They are difficult to show as being exhaustive or mutually exclusive. The list of possible causal agents is reasonably illustrative, but difficult to argue that it encompasses all agency possibilities (what about individual will, for example?); variance and process theories would seem to focus on entities in their permanence versus entities in the ways they change, but what about social forces (e.g. network theories); patterns of change (e.g. evolution, dialectics, cycles); complexity theories (e.g. non-deterministic systems), and the like. Moreover, it is not clear that these cannot be mixed and matched [see 16] by defining relationships among combinations of entities that are stable while others change. And finally level of analysis, given the uncountable ways to divide and subdivide hierarchies also seems like a malleable basis for a taxonomy.

Addressing the first issue of how the categories are constructed within Gregor (2006) is in itself interesting. The first step is identifying a dimension on which theories, one presumes instances of theory, can vary. The particular dimension she selects pertains to "goals" of which four (analysis and description, explanation, prediction, and prescription) are presented. In a footnote it is explained that these four goals derive from writing by Aristotle regarding four explanations of any "thing". Frankly, to our thinking, these are a reasonable and intuitive set of goals that would differentiate different types of theory. Interestingly, though a fifth category is derived from the combination of explanation and prediction. However, there is no explanation for why the other 9 logically possible combinations (5 more of pairs; 3 of sets of 3; and 1 of all four) are not equally proposed. Both "explanation and prescription" or "analysis and prediction" would seem likely to generate some instances. I "explain" that giving bonuses makes employees have higher morale, therefore for low morale I "prescribe" giving bonuses. Alternatively, I have "analyzed" that white collar workers have more autonomy (by observation or by definition), therefore I "predict" there will be more variance in how they make decisions or take actions (relative to blue collar workers). The reader can work out whether there are reasonable examples for the other combinations. That said, the fifth category of "explanation and prediction" seems perfectly reasonable. For my students (first author), I illustrate this category with the related theory that taking aspirin will, all else equal, bring down a fever (prediction) and, does this in part by suppressing the normal functioning of platelets, according to Wikipedia (explanation). Note how we could have only a predication or only an explanation in this case, but that the linkage of the two intuitively provides additional value. However, it is worth considering that much prediction, if it doesn't use established categories, may include an analysis or taxonomical component as a stated or unstated precursor to explanation or prediction. The point of this observation though is the probability that Gregor's categories of theory are NOT comprehensive. She acknowledges as much discussing the use of "critical theory" as possibly being considered a type of theory without fitting into any of the five categories but rather aiming at ways to improve people's lives.

Further, there is no clear demonstration that the categories are mutually exclusive. It is not clear why the same theory as a statement cannot be prescriptive and predictive? The wording might be a little unusual, such as you should give raises to productive employees because this will add to morale – which would seem to be both prescriptive and predictive. There is no category for both of these in the given typology.

What is clear from these rather pedantic criticisms is that the value of the Gregor categorization scheme is not dependent on being either comprehensive or mutually exclusive. It neither provides guidelines for assessing a new theory and assigning it to a category nor shows how a set of say 100 theories can be segregated by theory type with clear commonalities among category instances and differences across category instances. Which leaves the question of what, then, is the source of value? Doty and Glick [11] present the idea of ideal types which represent a sort of definitional type with or without any exactly matching instances (like an ideal circle is all points equidistant from a chosen point, but of course there are no instances because points are one dimensional and thus cannot be produced in a three dimensional world). This notion is also similar to the psychological concept of a prototype – a set of ideal attributes for an instance of a category [43]. In this sense Gregor's theory types are ideal types not meant to organize actual sets of theory instances, but rather to highlight differences on one possible dimension and elevate the importance of that dimension.

We do not argue against the wisdom of using goals as a basis for the taxonomy Gregor presents, however, we emphasize that the underlying question is not whether this categorization is true to the way reality is, but rather whether it is useful. In other words, a question of judgment, utility, and preference rather than truth and demonstrability. Those coming after Gregor are free to propose alternate formulations based on the Markus and Robey dimensions or whatever others they can conceive.

Note, though, that Gregor's dimensions can (1) take on a life of their own and come to be viewed as "the categories" rather than as one of a set of possible categories; (2) help guide and shape the thinking for followers who find them logical and "good enough" for their own uses; and (3) relieve others from the burden of having to define the categories anew for every further use.[3]

[3] Gregor (2006) adds an intriguing comment about grounded theory as theory emergent from a grounded process, but having read both process and outcome works of grounded theory it is not clear why the product of such work, if it stands independent of its specific environment, would not be amenable to being treated as analysis, prediction, explanation, or prescription). In other words, once formulated as a theory, would we know (or care) whether it came from a grounded theory process, another inductive process, a deductive process, or extrapolation from another field?

How should we understand "analysis" as theory?

The second area of examination of this paper pertains to her first theory type entitled "analysis" which highlights the description of a domain of interest and the development of a typology for that domain per se. Interestingly some other scholars, notably Weber [48] do not tend to recognize this category as a type of theory at all suggesting it may at best be a kind of "model" or pre-theoretic construct.

However, the discussion of the Gregor process for arriving at a set of categories is really an exercise in reflexivity, since the first of her categories itself pertains to categorization as a sort of theory. We wish to examine in more detail how the development and use of categorization can enhance research, both in terms of theory building and testing, and also more broadly in terms of knowledge accumulation.

This first of her categories of theory is called "analysis". It is distinguished in the definitional table (p. 62) not by what it is but by what it is not. It is theory that does not specify causal relationships (leaving open the question of whether it specifies other sorts of relationships like correlation, sequence, or probabilistic effects). It also is theory that does not specify prediction – and clearly this would conflict with the Popperian view of the sanctity of falsification in the formulation of theoretical statements as a way to distinguish between those that can and those that cannot be refuted (assuming the gathering of conflicting empirical evidence.)

The discussion emphasizes the use of typology and taxonomy as a central element in such analysis. The illustration of this sort of theory is based on a paper by Iivari, Hirschheim, and Klein [18] which presents a set of four IS development approaches pointing out "similarities and differences between them". (p. 623). Essentially, the four approaches represent a typology suggesting that each characterizes a different way to develop new systems. Presumably each development instance in practice would fall within one or another category, though these days we are seeing the rise of hybrid systems development approaches that combine elements of agile and traditional approaches. Projecting the likely intent of this category of theory is the notion that creating such a categorization scheme may provide significant value in and of itself in terms of its usefulness for understanding the domain. It may also provide input into other categories of theory such as predicting that agile methods will outperform (in cost, time, scope, quality or other measures) traditional methods at least relative to small customer facing applications. Note the symbiotic relationship between the content of the taxonomy as a necessary ingredient for developing the testable theoretical statements pertaining to performance of instances within each category. There is no logical prediction theory of this type without the distinct categories set forth in the taxonomy (or an alternative formulation). It is, of course, possible that the proponent of a testable theory of prediction may base statements on flimsy, ill-conceived and ad hoc categories (or strong ones) that have not been independently a source of reflection. Note further that the prediction test will have to assume (or measure) that each instance actually fits into that particular category in order to test contrasts between categories.

It is likely that the untestable nature of the development of categories is what leads some, notably Weber (2012) not to consider these as theories per se; but it is the utility of separate development and examination of these categories that in our view would lead to considering these as theory. It would seem that it takes little imagination to convert quality categorization schemes into prediction style theory by simply considering the characteristics of each category (e.g., instances in category 1 will be larger, more talented, quieter, etc. relative to those instances in category 2) or considering different productivity levels in using instances from each category.

Having accepted the assumption that "ad hoc categories" are fundamentally untestable with respect to the underlying reality (i.e., the instances they seek to organize), we can consider a different way of ascertaining their scientific utility. What researchers can (and should) test is the utility of categories for the consumer of these categories. In other words, when addressing the question of how "how good a categorization schema is?" one can test the relative benefit the user of the schema accrues as compared to something else. In this conception, categories are artifacts – figments of human imagination intended to serve as mental tools that help reduce complexity of reality. As tools, categories cannot be the sources of an objective truth; instead, they offer a pragmatic utility [40].

The tool perspective on categories leads us to formulate the following question: how can one establish the utility of categories? To address this question, it is important to note that categorization is a fundamental human mechanism. Psychologists argue that our access to reality is invariably mediated by categories – otherwise we would be incapable of dealing with the infinite diversity of the sensory experience - always unique and constantly changing [17, 43]. The implication of this for research in any given domain is that humans always operate on a set of explicit or implicit categories to make sense of this domain. In other words,

there is always a benchmark, a reference point, a status quo. A second consequence of the utilitarian, tool view of categories is the need for guidelines for how apply them unambiguously and consistently (i.e., having a manual to accompany the tool). Here, it is noteworthy to consider Gregor's (2006) prescription for formatting theories of the categorization, descriptive, or analytic type:

> "The logic for the placement of phenomena into categories should be clear, as should the characteristics that define each category. In addition important categories or elements should not be omitted from the classification system, that is, it should be complete and exhaustive (p. 624)."

We note the term "should" can be interpreted across a range from "must" to "would be nice if". We would be highly sympathetic to the latter but find the former interpretation to be unrealistic. To the extent that phenomena *DO* fit nicely into such categories or where a particular definition is clear and straightforward to apply, this is handy and helpful. Perhaps it is a goal to strive for as categories evolve. However, it is a tough criterion to pass when a field is opening to new and initial scrutiny. We suggest that the theory typology presented in this paper is extraordinarily useful, worthy of publication, but might not fit within this definition in its most rigorous interpretation. On the other hand, if a classification scheme, for example, rightly classifies a number of instances and defines a central tendency, even if not all instances, it can be of immense utility (as we argue Gregor's (2006) schema does.) Similarly, the criteria of completeness and mutual exclusivity, while a desirable target, are unlikely to be achieved and challenging if not impossible to demonstrate. If we needed to achieve these criteria as a "must" for categorization, we would likely have to do without categories at all – at least in our literature if not in our cognitive lives.

4 CONCLUSIONS

In this paper, we provided a more nuanced exposition of the nature of categories. Specifically, we show that fundamentally, two kinds of contributions can be made in research that seeks to propose new categories. First, researchers may identify natural discontinuities in reality, and propose categories to conceptualize these. If this is the intent of research, researchers need to be forthcoming and explicit in stating this objective. Having committed to this cause, researchers then need to demonstrate how the new categories make "deep" partitions of reality. This can be done by showing that the categories are internally strongly cohesive, but are also loosely coupled. In addition, the categories should offer rich inferences and capture bundles of inter-correlated attributes.

In contrast, other researchers may seek to propose "ad hoc categories". These categories are fundamentally utilitarian, and do not have to exhibit the properties of "common categories". When proposing a new classification schema based on these categories, researchers should be explicit in suggesting the added benefit of the proposed schema with the explicit or implicit schema or schemata that already exists. Such comparisons are most likely to be based on logical argument and presentation of inferred benefits. The generation of new, thought provoking, and potentially useful research questions would indicate utility of the categories in and of themselves. However, empirical data about such categorization based on interview, survey, "applicability check" [44] or observational evidence might also prove useful.

Assuming the demonstration of utility of new categories, any specification of specific and unambiguous rules for assigning individual cases to categories should be presented with humility as they will likely be subject to change over time. Consider the "black swan" and the inability to inductively "prove" exceptions just because none have been observed. In fact, the debate over specific instances and their assignment to categories can show differences in perceptions of that instance leading to greater understanding of particular cases (e.g. discovery or invention of new attributes) as well as the relative importance of identified attributes. To the extent that an abundance of instances exist, another approach to categories would be to demonstrate the population of each category (this is sometimes done with literature reviews assigning each underlying article into one or another category). Although once published such a sorting acquires a halo of acceptance, the discussion about these assignments and the nature of these instances or articles holds the potential for richer and deeper interpretations and elicitation of useful meaning.

Following this logic of fluid or ad hoc categories, future researchers using these categories may (1) find them applicable and create further inference based on them; (2) find "fuzziness" in their boundaries and further refine our understanding of them; (3) discover significant cases (perhaps presenting as "outliers") that populate the ubiquitous "other" category until their volume suggests a new substantial category; or (4) simply present new

conundrums and raise innovative new problems or questions.

To be more specific, we see applications for examination of categories throughout the array of topics of relevance to those studying "computer personnel" or "computation and people". For example, recent scientific research calls into question the simplistic categorization of all people by either sex or gender as "male" or "female" without consideration of the many potential variations on the levels of both chromosomes and social roles. Much management and MIS literature particularly pertaining to social inclusion tends to be based on old and often unquestioned categories that do not necessarily take into account shades and variations of biological and cultural distinctions.

Another example pertains to IT workers. Recent literature has discussed the decades long movement of the use of computing from the exclusive realm of IT professionals to just about every worker from utility meter readers to robotics specialists directing mechanisms to pour molten metals into molds [32]. Perhaps it is time to rethink meaningful ways to consider and treat differentiations among those who work with computing in different ways that would be meaningful for understanding the evolution of these people and jobs. Such categorization has implications for approaches to train, hire, retain, and make use of their labors.

Thus, we encourage researchers to embrace a more fluid, more fleeting conceptualization of categories as elements of research contribution. As we demonstrated on the analysis of a seminal categorization framework by Gregor (2006), a given set of categories can be criticized and challenged, as many alternative ways to organize phenomena in a domain are possible – even while its initial "common category" presentation also demonstrates usefulness. We suggest for researchers to demonstrate the utility of categories to their consumers, and systematically and rigorously compare the new categories proposed to those already in existence. At the same time, we recommend for reviewers and editors to (1) treat categories presented as if they were solid and permanent with some skepticism while seemingly paradoxically (2) being open to varied and diverse approaches to categorization within any given domain – especially ones where categories have been accepted and untested for a long period of time. We believe that IS research can meaningfully increase its accumulated knowledge base by considering the character of the categories and instances as well as the relationships between them within IS overall as well as each sub-discipline

REFERENCES

[1] Barsalou, L.W. 1983. Ad hoc categories. *Memory & Cognition*. 11, (1983), 211–227.

[2] Barsalou, L.W. et al. 1998. Basing categorization on individuals and events. *Cognitive psychology*. 36, 3 (1998), 203–272.

[3] Benbasat, I. and Barki, H. 2007. Quo vadis TAM? *Journal of the association for information systems*. 8, 4 (2007), 7.

[4] Bergholtz, M. and Eriksson, O. 2015. Towards a Socio-Institutional Ontology for Conceptual Modelling of Information Systems. *Advances in Conceptual Modeling*. Springer. 225–235.

[5] Berlin, B. et al. 1966. Folk Taxonomies and Biological Classification. *Science*. 154, 3746 (1966), 273–275.

[6] Borges, J.L. 1966. *Other inquisitions, 1937-1952*. Washington Square Press,.

[7] Bowker, G.C. and Star, S.L. 2000. *Sorting things out: classification and its consequences*. MIT Press.

[8] Bunge, M. 1977. *Treatise on basic philosophy: Ontology I: the furniture of the world*. Reidel.

[9] Carey, S. 2009. *The Origin of Concepts*. Oxford University Press.

[10] Castellanos, A. et al. 2016. Conceptual Modeling in Open Information Environments. *AIS SIGSAND Symposium* (Lubbock, TX, 2016), 1–7.

[11] Doty, D.H. and Glick, W.H. 1994. Typologies as a unique form of theory building: Toward improved understanding and modeling. *Academy of management review*. 19, 2 (1994), 230–251.

[12] Emirbayer, M. 1997. Manifesto for a relational sociology. *American journal of sociology*. 103, 2 (1997), 281–317.

[13] Eriksson, O. et al. 2018. Institutional ontology for conceptual modeling. *Journal of Information Technology*. (2018), 1–19.

[14] Fiss, P.C. 2011. Building better causal theories: A fuzzy set approach to typologies in organization research. *Academy of Management Journal*. 54, 2 (2011), 393–420.

[15] Gregor, S. 2006. The nature of theory in information systems. *MIS Quarterly*. 30, 3 (Sep. 2006), 611–642.

[16] de Guinea, A.O. and Webster, J. 2017. Combining variance and process in information systems research: Hybrid approaches. *Information and Organization*. 27, 3 (2017), 144–162.

[17] Harnad, S. 2005. To Cognize is to Categorize: Cognition is Categorization. H. Cohen and C. Lefebvre, eds. Elsevier Science. 20–45.

[18] Iivari, J. et al. 2000. A dynamic framework for classifying information systems development methodologies and approaches. *Journal of Management Information Systems*. 17, 3 (2000), 179–218.

[19] Lakoff, G. 1987. *Women, fire, and dangerous things : what categories reveal about the mind*. University of Chicago Press.

[20] Larsen, K. and Bong, C.H. 2016. A Tool for Addressing Construct Identity in Literature Reviews and Meta-Analyses. *MIS Quarterly*. 40, 3 (2016), 1–23.

[21] Lassaline, M.E. et al. 1992. Basic Levels in Artificial and Natural Categories: Are All Basic Levels Created Equal? *Percepts, Concepts and Categories: The Representation and Processing of Information*. B. Barbara, ed. North-Holland. 328–378.

[22] Lukyanenko, R. et al. 2016. Emerging problems of data quality in citizen science. *Conservation Biology*. 30, 3 (2016), 447–449.

[23] Lukyanenko, R. et al. 2016. Participatory Design for User-generated Content: Understanding the challenges and moving forward. *Scandinavian Journal of Information Systems*. 28, 1 (2016), 37–70.

[24] Lukyanenko, R. et al. 2017. Representing Crowd Knowledge: Guidelines for Conceptual Modeling of User-generated Content. *Journal of the Association for Information Systems*. 18, 4 (2017), 297–339.

[25] Lukyanenko, R. et al. 2014. The IQ of the Crowd: Understanding and Improving Information Quality in Structured User-generated Content. *Information Systems Research*. 25, 4 (2014), 669–689.

[26] Lukyanenko, R. and Samuel, B.M. 2017. Are all Classes Created Equal? Increasing Precision of Conceptual Modeling Grammars. *ACM Transactions on Management Information Systems (TMIS)*. 40, 2 (Forthcoming 2017), 1–25.

[27] March, S. and Allen, G. 2015. Classification with a purpose. *Symposium on Research in Systems Analysis and Design* (Richmond, VA, 2015), 1–10.

[28] March, S.T. and Allen, G.N. 2014. Toward a social ontology for conceptual modeling. *Communications of the AIS*. 34, (2014).

[29] Markus, M.L. and Robey, D. 1988. Information technology and organizational change: causal structure in theory and research. *Management science*. 34, 5 (1988), 583–598.

[30] Medin, D.L. et al. 2000. Are there kinds of concepts? *Annual review of psychology*. 51, 1 (2000), 121–147.

[31] Murphy, G. 2004. *The big book of concepts*. MIT Press.

[32] Niederman, F. et al. 2016. On the co-evolution of information technology and information systems personnel. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*. 47, 1 (2016), 29–50.

[33] Ogunseye, S. et al. 2017. Do Crowds Go Stale? Exploring the Effects of Crowd Reuse on Data Diversity. *WITS 2017* (Seoul, South Korea, 2017).

[34] Parsons, J. and Wand, Y. 2008. A question of class. *Nature*. 455, 7216 (Oct. 2008), 1040–1041.

[35] Parsons, J. and Wand, Y. 2013. Cognitive Principles to Support Information Requirements Agility. *Advanced Information Systems Engineering Workshops* (2013), 192–197.

[36] Parsons, J. and Wand, Y. 2000. Emancipating Instances from the Tyranny of Classes in Information Modeling. *ACM Transactions on Database Systems*. 25, 2 (2000), 228–268.

[37] Parsons, J. and Wand, Y. 2012. Extending Classification Principles from Information Modeling to Other Disciplines. *Journal of the Association for Information Systems*. 14, 5 (2012), 2.

[38] Parsons, J. and Wand, Y. 2008. Using cognitive principles to guide classification in information systems modeling. *MIS Quarterly*. 32, 4 (Dec. 2008), 839–868.

[39] Pentland, B. et al. 2017. Capturing reality in flight? Empirical tools for strong process theory. *ICIS 2017* (Seoul, South Korea, 2017), 1–12.

[40] Purao, S. 2013. Truth or dare: The ontology question in design science research. *Journal of Database Management (JDM)*. 24, 3 (2013), 51–66.

[41] Ritvo, H. 1997. *The platypus and the mermaid, and other figments of the classifying imagination*. Harvard University Press.

[42] Rosch, E. et al. 1976. Basic Objects in Natural Categories. *Cognitive Psychology*. 8, 3 (1976), 382–439.

[43] Rosch, E. 1978. Principles of Categorization. E. Rosch and B. Lloyd, eds. John Wiley & Sons Inc. 27–48.

[44] Rosemann, M. and Vessey, I. 2008. Toward improving the relevance of information systems research to practice: the role of applicability checks. *MIS Quarterly*. 32, 1 (2008), 1–22.

[45] Searle, J.R. 1995. *The construction of social reality*. Simon and Schuster.

[46] Sidorova, A. et al. 2008. Uncovering the intellectual core of the information systems discipline. *MIS Quarterly*. (2008), 467–482.

[47] Smith, L.B. 2005. Emerging ideas about categories. L. Gershkoff-Stowe and D.H. Rakison, eds. L. Erlbaum Associates. 159–175.

[48] Weber, R. 2012. Evaluating and Developing Theories in the Information Systems Discipline. *Journal of the Association for Information Systems*. 13, 1 (2012), 1–30.